

RECEIVED

1645

AUG 09 2002

TECH CENTER 1600/2900



ENTERED 1600

P#8

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/674,183DATE: 08/01/2002
TIME: 11:19:56Input Set : A:\seqlist.txt
Output Set: N:\CRF3\08012002\I674183.raw

3 <110> APPLICANT: RAPPUOLI, Rino
 4 GRANDI, Guido
 6 <120> TITLE OF INVENTION: POLYPEPTIDE CARRIER PROTEIN
 8 <130> FILE REFERENCE: PP00362.102 / 2302-0362
 10 <140> CURRENT APPLICATION NUMBER: 09/674,183
 11 <141> CURRENT FILING DATE: 2000-11-14
 13 <150> PRIOR APPLICATION NUMBER: PCT/IB99/00844
 14 <151> PRIOR FILING DATE: 1999-04-27
 16 <150> PRIOR APPLICATION NUMBER: GB9808932.9
 17 <151> PRIOR FILING DATE: 1998-04-27
 19 <160> NUMBER OF SEQ ID NOS: 20
 21 <170> SOFTWARE: PatentIn Ver. 2.1
 23 <210> SEQ ID NO: 1
 24 <211> LENGTH: 15
 25 <212> TYPE: PRT
 26 <213> ORGANISM: T-cell epitope P23TT
 28 <400> SEQUENCE: 1
 29 Val Ser Ile Asp Lys Phe Arg Ile Phe Cys Lys Ala Asn Pro Lys
 30 1 5 10 15
 33 <210> SEQ ID NO: 2
 34 <211> LENGTH: 16
 35 <212> TYPE: PRT
 36 <213> ORGANISM: T-cell epitope P32TT
 38 <400> SEQUENCE: 2
 39 Leu Lys Phe Ile Ile Lys Arg Tyr Thr Pro Asn Asn Glu Ile Asp Ser
 40 1 5 10 15
 43 <210> SEQ ID NO: 3
 44 <211> LENGTH: 16
 45 <212> TYPE: PRT
 46 <213> ORGANISM: T-cell epitope P21TT
 48 <400> SEQUENCE: 3
 49 Ile Arg Glu Asp Asn Asn Ile Thr Leu Lys Leu Asp Arg Cys Asn Asn
 50 1 5 10 15
 53 <210> SEQ ID NO: 4
 54 <211> LENGTH: 18
 55 <212> TYPE: PRT
 56 <213> ORGANISM: T-cell epitope PF T3
 58 <400> SEQUENCE: 4
 59 Glu Lys Lys Ile Ala Lys Met Glu Lys Ala Ser Ser Val Phe Asn Val
 60 1 5 10 15
 62 Val Asn
 66 <210> SEQ ID NO: 5
 67 <211> LENGTH: 21

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/674,183

DATE: 08/01/2002
TIME: 11:19:56

Input Set : A:\seqlist.txt
Output Set: N:\CRF3\08012002\I674183.raw

68 <212> TYPE: PRT
69 <213> ORGANISM: T-cell epitope P30TT
71 <400> SEQUENCE: 5
72 Phe Asn Asn Phe Thr Val Ser Phe Trp Leu Arg Val Pro Lys Val Ser
73 1 5 10 15
75 Ala Ser His Leu Glu
76 20
79 <210> SEQ ID NO: 6
80 <211> LENGTH: 14
81 <212> TYPE: PRT
82 <213> ORGANISM: T-cell epitope P2TT
84 <400> SEQUENCE: 6
85 Gln Tyr Ile Lys Ala Asn Ser Lys Phe Ile Gly Ile Thr Glu
86 1 5 10
89 <210> SEQ ID NO: 7
90 <211> LENGTH: 13
91 <212> TYPE: PRT
92 <213> ORGANISM: T-cell epitope HA
94 <400> SEQUENCE: 7
95 Pro Lys Tyr Val Lys Gln Asn Thr Leu Lys Leu Ala Thr
96 1 5 10
99 <210> SEQ ID NO: 8
100 <211> LENGTH: 20
101 <212> TYPE: PRT
102 <213> ORGANISM: T-cell epitope HBVnc
104 <400> SEQUENCE: 8
105 Pro His His Thr Ala Leu Arg Gln Ala Ile Leu Cys Trp Gly Glu Leu
106 1 5 10 15
108 Met Thr Leu Ala
109 20
112 <210> SEQ ID NO: 9
113 <211> LENGTH: 15
114 <212> TYPE: PRT
115 <213> ORGANISM: T-cell epitope HBsAg
117 <400> SEQUENCE: 9
118 Phe Phe Leu Leu Thr Arg Ile Leu Thr Ile Pro Gln Ser Leu Asp
119 1 5 10 15
122 <210> SEQ ID NO: 10
123 <211> LENGTH: 16
124 <212> TYPE: PRT
125 <213> ORGANISM: T-cell epitope MT
127 <400> SEQUENCE: 10
128 Tyr Ser Gly Pro Leu Lys Ala Glu Ile Ala Gln Arg Leu Glu Asp Val
129 1 5 10 15
132 <210> SEQ ID NO: 11
133 <211> LENGTH: 20
134 <212> TYPE: PRT
135 <213> ORGANISM: T-cell epitope HSP70
137 <400> SEQUENCE: 11

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/674,183

DATE: 08/01/2002
TIME: 11:19:56

Input Set : A:\seqlist.txt
Output Set: N:\CRF3\08012002\I674183.raw

```

138 Gln Pro Ser Val Gln Ile Gln Val Tyr Gln Gly Glu Arg Glu Ile Ala
139   1           5           10          15
141 Ser His Asn Lys
142       20
145 <210> SEQ ID NO: 12
146 <211> LENGTH: 8
147 <212> TYPE: PRT
148 <213> ORGANISM: T-cell epitope Flag peptide
150 <400> SEQUENCE: 12
151 Met Asp Tyr Lys Asp Asp Asp Asp
152   1           5
155 <210> SEQ ID NO: 13
156 <211> LENGTH: 657
157 <212> TYPE: DNA
158 <213> ORGANISM: Recombinant N10 construct
160 <220> FEATURE:
161 <221> NAME/KEY: CDS
162 <222> LOCATION: (1)..(657)
164 <400> SEQUENCE: 13
165 atg ggg ggt tct cat cat cat cat cat ggt atg gct agc atg gat      48
166 Met Gly Gly Ser His His His His His Gly Met Ala Ser Met Asp
167   1           5           10          15
169 tac aag gac gac gat gat atc gaa ggt cgc aaa ggt gtt tcc atc gac      96
170 Tyr Lys Asp Asp Asp Asp Ile Glu Gly Arg Lys Gly Val Ser Ile Asp
171       20          25          30
173 aaa ttc cgt atc ttc tgc aaa gct aac ccg aaa aaa ggt ctg aaa ttc      144
174 Lys Phe Arg Ile Phe Cys Lys Ala Asn Pro Lys Lys Gly Leu Lys Phe
175       35          40          45
177 atc atc aaa cgt tac acc ccg aac aac gaa atc gac tcc aaa ggt atc      192
178 Ile Ile Lys Arg Tyr Thr Pro Asn Asn Glu Ile Asp Ser Lys Gly Ile
179       50          55          60
181 cgt gaa gac aac aac atc acc ctg aaa ctg gac cgt tgc aac aac aaa      240
182 Arg Glu Asp Asn Asn Ile Thr Leu Lys Leu Asp Arg Cys Asn Asn Lys
183       65          70          75          80
185 ggt gaa aag aag atc gct aaa atg gaa aaa gct tct tct gtt ttc aac      288
186 Gly Glu Lys Lys Ile Ala Lys Met Glu Lys Ala Ser Ser Val Phe Asn
187       85          90          95
189 gtt gtt aac tct aaa ggt ttc aac aac ttc acc gtt tcc ttc tgg ctg      336
190 Val Val Asn Ser Lys Gly Phe Asn Asn Phe Thr Val Ser Phe Trp Leu
191       100         105         110
193 cgt gtt ccg aaa gtt tcc gct tcc cac ctg gaa aaa ggt cag tac atc      384
194 Arg Val Pro Lys Val Ser Ala Ser His Leu Glu Lys Gly Gln Tyr Ile
195       115         120         125
197 aaa gct aac tcc aaa ttc atc ggt atc acc gaa aaa ggt gga tct ccg      432
198 Lys Ala Asn Ser Lys Phe Ile Gly Ile Thr Glu Lys Gly Gly Ser Pro
199       130         135         140
201 cat cat acc gcg ctg cgc cag gcg att ctg tgc tgg ggc gaa ctg atg      480
202 His His Thr Ala Leu Arg Gln Ala Ile Leu Cys Trp Gly Glu Leu Met
203       145         150         155         160

```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/674,183

DATE: 08/01/2002
TIME: 11:19:56

Input Set : A:\seqlist.txt
Output Set: N:\CRF3\08012002\I674183.raw

```

205 acc ctg gcg aaa gga tct ccg aaa tat gtg aaa cag aac acc ctg aaa 528
206 Thr Leu Ala Lys Gly Ser Pro Lys Tyr Val Lys Gln Asn Thr Leu Lys
207 165 170 175
209 ctg gcg acc aaa gga tcg ttt ttt ctg ctg acc cgc att ctg acc att 576
210 Leu Ala Thr Lys Gly Ser Phe Phe Leu Leu Thr Arg Ile Leu Thr Ile
211 180 185 190
213 ccg cag tct ctg gat aaa ggc tat tct ggc ccg ctg aaa gcg gaa att 624
214 Pro Gln Ser Leu Asp Lys Gly Tyr Ser Gly Pro Leu Lys Ala Glu Ile
215 195 200 205
217 gcg cag cgc ctg gaa gat gtg aaa gga tcc taa 657
218 Ala Gln Arg Leu Glu Asp Val Lys Gly Ser
219 210 215
222 <210> SEQ ID NO: 14
223 <211> LENGTH: 218
224 <212> TYPE: PRT
225 <213> ORGANISM: Recombinant N10 construct
227 <400> SEQUENCE: 14
228 Met Gly Gly Ser His His His His His Gly Met Ala Ser Met Asp
229 1 5 10 15
231 Tyr Lys Asp Asp Asp Asp Ile Glu Gly Arg Lys Gly Val Ser Ile Asp
232 20 25 30
234 Lys Phe Arg Ile Phe Cys Lys Ala Asn Pro Lys Lys Gly Leu Lys Phe
235 35 40 45
237 Ile Ile Lys Arg Tyr Thr Pro Asn Asn Glu Ile Asp Ser Lys Gly Ile
238 50 55 60
240 Arg Glu Asp Asn Asn Ile Thr Leu Lys Leu Asp Arg Cys Asn Asn Lys
241 65 70 75 80
243 Gly Glu Lys Lys Ile Ala Lys Met Glu Lys Ala Ser Ser Val Phe Asn
244 85 90 95
246 Val Val Asn Ser Lys Gly Phe Asn Asn Phe Thr Val Ser Phe Trp Leu
247 100 105 110
249 Arg Val Pro Lys Val Ser Ala Ser His Leu Glu Lys Gly Gln Tyr Ile
250 115 120 125
252 Lys Ala Asn Ser Lys Phe Ile Gly Ile Thr Glu Lys Gly Gly Ser Pro
253 130 135 140
255 His His Thr Ala Leu Arg Gln Ala Ile Leu Cys Trp Gly Glu Leu Met
256 145 150 155 160
258 Thr Leu Ala Lys Gly Ser Pro Lys Tyr Val Lys Gln Asn Thr Leu Lys
259 165 170 175
261 Leu Ala Thr Lys Gly Ser Phe Phe Leu Leu Thr Arg Ile Leu Thr Ile
262 180 185 190
264 Pro Gln Ser Leu Asp Lys Gly Tyr Ser Gly Pro Leu Lys Ala Glu Ile
265 195 200 205
267 Ala Gln Arg Leu Glu Asp Val Lys Gly Ser
268 210 215
271 <210> SEQ ID NO: 15
272 <211> LENGTH: 432
273 <212> TYPE: DNA
274 <213> ORGANISM: Recombinant N6 construct

```

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/674,183

DATE: 08/01/2002
TIME: 11:19:56

Input Set : A:\seqlist.txt
Output Set: N:\CRF3\08012002\I674183.raw

276 <220> FEATURE:
 277 <221> NAME/KEY: CDS
 278 <222> LOCATION: (1)...(432)
 280 <400> SEQUENCE: 15
 281 atg ggg ggt tct cat cat cat cat cat ggt atg gct agc atg gat 48
 282 Met Gly Gly Ser His His His His His Gly Met Ala Ser Met Asp
 283 1 5 10 15
 285 tac aag gac gac gat gat atc gaa ggt cgc aaa ggt gtt tcc atc gac 96
 286 Tyr Lys Asp Asp Asp Asp Ile Glu Gly Arg Lys Gly Val Ser Ile Asp
 287 20 25 30
 289 aaa ttc cgt atc ttc tgc aaa gct aac ccg aaa aaa ggt ctg aaa ttc 144
 290 Lys Phe Arg Ile Phe Cys Lys Ala Asn Pro Lys Lys Gly Leu Lys Phe
 291 35 40 45
 293 atc atc aaa cgt tac acc ccg aac aac gaa atc gac tcc aaa ggt atc 192
 294 Ile Ile Lys Arg Tyr Thr Pro Asn Asn Glu Ile Asp Ser Lys Gly Ile
 295 50 55 60
 297 cgt gaa gac aac aac atc acc ctg aaa ctg gac cgt tgc aac aac aaa 240
 298 Arg Glu Asp Asn Asn Ile Thr Leu Lys Leu Asp Arg Cys Asn Asn Lys
 299 65 70 75 80
 301 ggt gaa aag aag atc gct aaa atg gaa aaa gct tct tct gtt ttc aac 288
 302 Gly Glu Lys Lys Ile Ala Lys Met Glu Lys Ala Ser Ser Val Phe Asn
 303 85 90 95
 305 gtt gtt aac tct aaa ggt ttc aac aac ttc acc gtt tcc ttc tgg ctg 336
 306 Val Val Asn Ser Lys Gly Phe Asn Asn Phe Thr Val Ser Phe Trp Leu
 307 100 105 110
 309 cgt gtt ccg aaa gtt tcc gct tcc cac ctg gaa aaa ggt cag tac atc 384
 310 Arg Val Pro Lys Val Ser Ala Ser His Leu Glu Lys Gly Gln Tyr Ile
 311 115 120 125
 313 aaa gct aac tcc aaa ttc atc ggt atc acc gaa aaa ggt gga tcc taa 432
 314 Lys Ala Asn Ser Lys Phe Ile Gly Ile Thr Glu Lys Gly Gly Ser
 315 130 135 140
 318 <210> SEQ ID NO: 16
 319 <211> LENGTH: 143
 320 <212> TYPE: PRT
 321 <213> ORGANISM: Recombinant N6 construct
 323 <400> SEQUENCE: 16
 324 Met Gly Gly Ser His His His His His Gly Met Ala Ser Met Asp
 325 1 5 10 15
 327 Tyr Lys Asp Asp Asp Asp Ile Glu Gly Arg Lys Gly Val Ser Ile Asp
 328 20 25 30
 330 Lys Phe Arg Ile Phe Cys Lys Ala Asn Pro Lys Lys Gly Leu Lys Phe
 331 35 40 45
 333 Ile Ile Lys Arg Tyr Thr Pro Asn Asn Glu Ile Asp Ser Lys Gly Ile
 334 50 55 60
 336 Arg Glu Asp Asn Asn Ile Thr Leu Lys Leu Asp Arg Cys Asn Asn Lys
 337 65 70 75 80
 339 Gly Glu Lys Lys Ile Ala Lys Met Glu Lys Ala Ser Ser Val Phe Asn
 340 85 90 95
 342 Val Val Asn Ser Lys Gly Phe Asn Asn Phe Thr Val Ser Phe Trp Leu

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/674,183

DATE: 08/01/2002

TIME: 11:19:57

Input Set : A:\seqlist.txt

Output Set: N:\CRF3\08012002\I674183.raw